A Panel Analysis of Polish Regional Cities Residential Price Convergence in the Primary Market

Matysiak, George and Olszewski, Krzysztof

Krakow University of Economics, Narodowy Bank Polski

23 June 2019

A Panel Analysis of Polish Regional Cities Residential Price Convergence in the Primary Market

George A. Matysiak*

Krakow University of Economics

Krzysztof Olszewski

Narodowy Bank Polski, Warsaw School of Economics

June 2019

Keywords: Polish regional residential prices, Phillips and Sul, panel convergence, relative transition, ordered logit model.

* Corresponding author: george.matysiak@uek.krakow.pl or gmatysiak@gmail.com

This article presents the opinions of its authors and not necessarily the official position of the Narodowy Bank Polski.

Abstract

We employ two methodologies in order to identify groupings of cities and to analyse the factors which drive convergence in residential prices across Polish prime markets over the period 2007-2018. The Phillips and Sul (2007) methodology is used to identify convergence in primary residential prices in the major Polish cities. The results indicate that residential prices do not converge to a single common trend. However, we find the existence of three distinct sub-groups of cities ('clubs') where residential prices converge to their own steady-state paths. Using an ordered logit model, we investigate supply and demand factors determining club membership, which subsequently allocates 13 out of the 15 cities as belonging to the clubs identified by the Phillips and Sul procedure.

Keywords

Polish residential prices, Phillips and Sul, panel convergence, relative transition, ordered logit model.

1. Introduction

The purpose of the research is to investigate the evolution of new residential property prices in sixteen major cities in the Polish market over the period 2007:Q1-2018:Q4. We focus on the market for newly constructed housing units, referred to as the "primary market" in the text. Whilst residential price differentials exist between cities and evolution of prices has been heterogeneous, there has been a narrowing of the price gap between locations over the twelve-year period. In this regard, we examine the evidence for convergence in residential prices in the Polish cities employing a methodology proposed by Phillips and Sul (2007). The approach also enables the identification of potential convergence 'clubs', where groups of cities have a tendency to converge to similar price levels. As evidence is found for such clubs we then investigate the common factors which explain the formation of the clubs. Such knowledge can be valuable for housing policymakers as well as for financial stability purposes.

The paper is organised as follows. Section 2 provides an overview of Polish residential markets. Section 3 presents a brief literature review. Section 4 outlines the convergence methodology we employ. Section 5 reports the results of the convergence analysis. Section 6 outlines the ordered logit model methodology used to analyse the clubs and reports the results. Section 7 concludes with observations and conclusions.

2. Overview of Polish residential markets

Poland, similar to other post-socialist countries, had a significant lack of housing when it entered the EU in 2004. This was observed in three dimensions namely, the number of dwellings per 1000 of population, size of the average flat and finally, the quality of accommodation. According to Eurostat (2018) data, most of the Central and Eastern European countries in 2016 had an overcrowding rate in excess of 32%. By way of comparison, the corresponding figure was less than 8% in Germany, France and Spain. The average number of rooms per person was one for the CEE countries, being in excess of one-and-a-half for older member states.

It is difficult to measure quality objectively, however, it is common knowledge that the majority of housing in Poland that was built during the 1970s and 1980s used the prefabricated technology and low-quality materials. These factors prompted potential buyers to demand new housing from the primary market, making such housing affordable given the improving economic backdrop.

Ciarlone (2015) points out that the price increases in Central Eastern European countries were caused by an insufficient number of available flats. As incomes grew, people could afford housing and, as the demand for housing increased, property prices rose. The empirical analysis of the primary housing market in Warsaw by Augustyniak et al. (2014) shows that housing demand reacted strongly to interest rates and the availability of mortgages.

Leszczyński and Olszewski (2017), using annual data, analysed house prices in the 17 largest cities in Poland, for both the primary and the secondary markets. They found that house prices increase with growing wages, construction costs, average value of mortgages and decline with higher unemployment and real interest rates. The values of the respective parameters in each market are such that they are significantly larger for the secondary market (existing stock) than for the primary market. It may be the case that more affluent people purchasing on the primary market are less financially constrained than those in the secondary market. A further contributing factor may be that the proceeds from the sale of old property are used to finance the difference between the values of the new and old house. This process is known as ‘filtration’, where an old property gets filtered down to less affluent people.

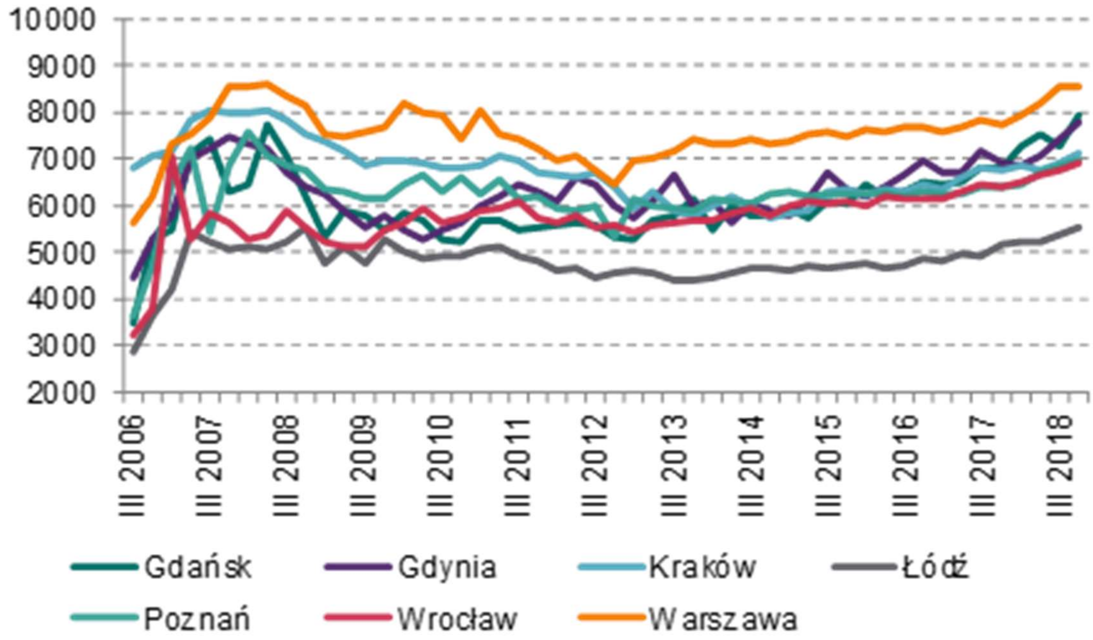
When Poland entered the EU, mortgages by banks started to grow at a fast pace, especially mortgages in foreign currencies. As foreign currency mortgages, particularly those denominated in Swiss Francs, carried lower interest rates, more expensive properties could be purchased for the same monthly instalment compared to purchases employing local currency loans. This factor not only contributed towards house price growth, but also generated systemic risk when the global financial crisis broke out. Fortunately, this risk did not materialize and the foreign currency mortgages are continuously repaid.

The Polish exchange rate vis-a-vis the Swiss Franc and the Euro fell, from its high value relative to these currencies in 2008, by around 50% in 2009. At the time some 40 per cent of Polish home loans were denominated in Swiss Francs. Fortunately, only adjustable interest mortgages were and are still offered in Poland. Shortly after, in order to support its own exchange rate and the European economy, both the Swiss National Bank and the ECB decreased interest rates. As a consequence, the increase in monthly repayment instalments due to the exchange rate shock in 2009 was partly offset by the decline of interest rates. However, the impact of the fall in the value of the Polish currency meant that the value of Swiss Franc denominated mortgages increased in Polish Zloty[1].

Broadly, residential property prices rose sharply just after the Polish EU accession in May 2004, and started to decline when the global financial crisis took hold. From around 2012 prices began to rise again. Figure 1 shows the prices for the 7 largest cities over the period 2006:Q3 to 2018:Q3.

[1] A detailed case study is presented by Łaszek et al. (2016a).

Figure 1: Transaction prices per square metre of housing in the largest 7 cities in the primary market



Source: NBP (2019).

We emphasise that price developments depend not only on the interaction between supply and demand, but also on the price setting decisions of developers. Łaszek et al. (2016b) analysed individual transaction prices in the primary market in Warsaw and, employing a hedonic regression model based on property fundamentals, compared them to the implied hedonic value. The factual prices were distributed around the hedonic value but with significant observed dispersion, with properties being sold above and below their hedonic value. This pattern of dispersion varied over the period 2006-2014. The better the market conditions were the more developers could differentiate sales prices, without bearing the risk of being stuck with unsold properties.

2.1 Profile of residential prices

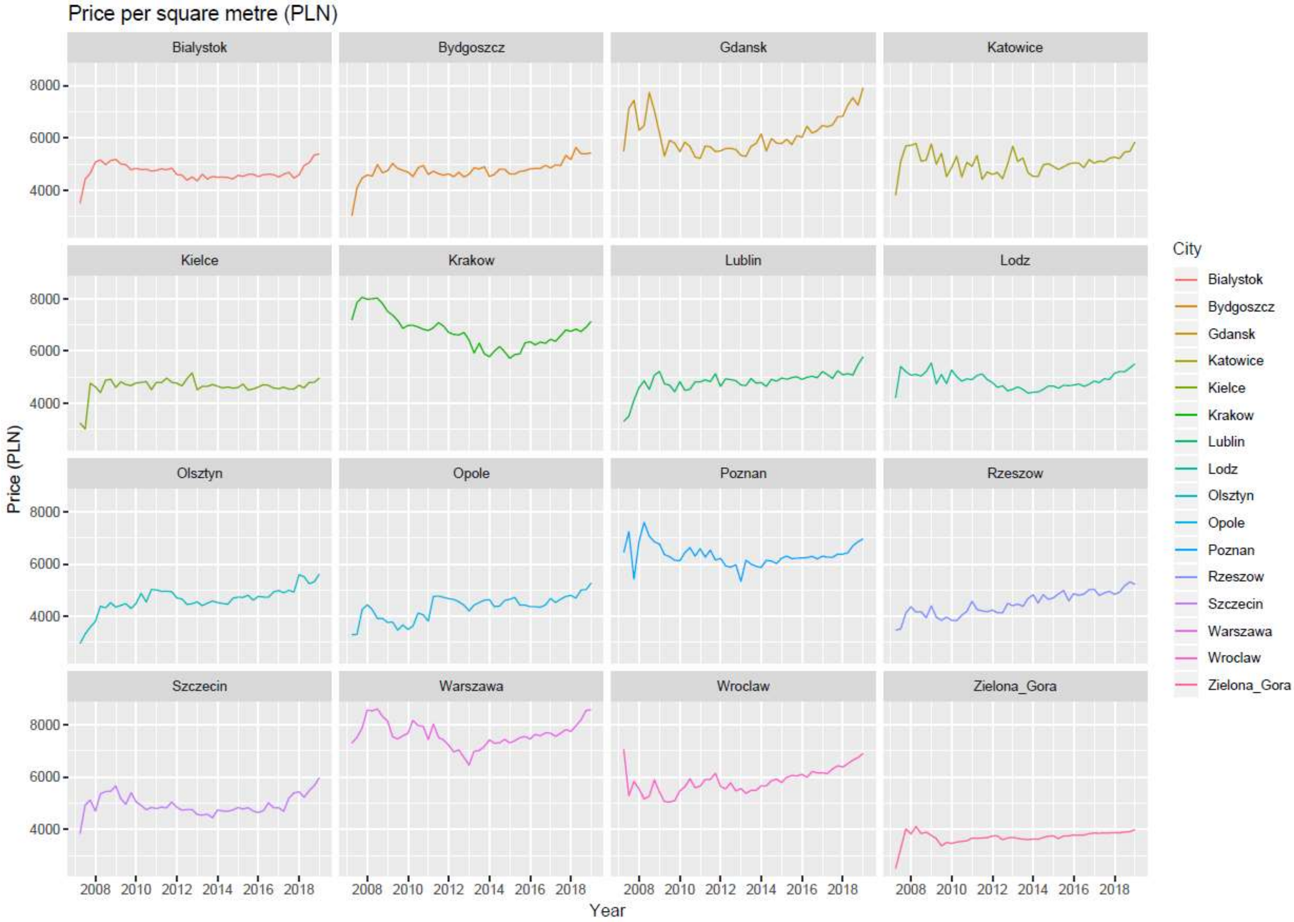
The housing markets in the 16 voivodeship capital cities which we investigate are differentiated in terms of size and in terms of price levels. Baldowska et al. (2014) applied a cluster analysis to determine static clusters for the year 2012, finding different clusters when considering a variety of variables. The authors start with an obvious grouping by city size, which allows them to identify two groups of cities, namely 7 cities with more than 400 thousand inhabitants and 9 cities with smaller populations[2]. The number of clusters and their composition differs considerably when other variables such as the housing situation, demographics or economic conditions are taken into account. Consequently, we take the view that groupings of cities should take into account the evolving profile of residential prices over time, including analysis of the supply and demand factors contributing towards price convergence.

Figure 2 shows the profile of residential prices for the 16 voivodeship capital cities over the period 2007:Q1-2018:Q4. A number of broad patterns can be observed. Although fluctuations in prices are clearly visible, it is seen that the overall trend for most cities over the second half of the period is largely one of rising prices, being more pronounced for the largest cities.

Demand in the larger cities has been driven, in addition to local demand, by individuals from other locations, partly for rental purposes and partly in the expectation that the local economy will grow and that house prices would follow. This was especially visible in the years 2006–2007 and lead to excessive prices, which subsequently declined in may locations around 2012. When the perception of the whole economy was particularly positive, the largest cities again started to attract migrations, together with resumed purchases by individuals from other locations. This explains why the significant difference in house price profiles can be found between the larger and the smaller cities. The largest cities experienced the previously described U-shaped type price development, whilst the group of smaller cities displayed a slowly growing trend in property prices, as observed for Bydgoszcz, Katowice, Lublin, Olsztyn, Opole, Rzeszów and Zielona Góra. This was the result of the slow, but continuous, increase in demand for housing in these smaller cities. However, notable exceptions are Białystok, and Kielce, where prices had remained relatively flat over most of the period.

[2] For example, the first cluster, by size, consists of Gdańsk, Cracow, Łódź, Poznań, Szczecin, Warsaw and Wrocław.

Figure 2: Distribution of prices 2007:Q1-2018:Q4



Source: NBP (2019) data.

3. Literature review

The concept of convergence is well developed and has an extensive theoretical and applied research literature. Convergence processes and the underlying dynamics have their roots in the economics literature, and have been extensively addressed in the context of economic growth, especially real per capita incomes across regions and countries. Frequently cited influential early works include: Baumol (1986), Barro (1991), Mankiw et al (1992) and Barro and Sala-i-Martin (1992). Much of the early literature employed versions of the Solow growth model Solow (1956) and the concept of Beta-convergence (β-convergence), which continues to be largely used in the growth literature.

Several definitions of 'convergence', for example, Beta (β-convergence) and sigma (σ-convergence), have emerged and been applied in empirical studies in many areas. The former measures the process by which poorer economies grow faster than wealthier economies, and subsequently catch-up (converge) with the wealthy economies. The σ-convergence measures the reduction in the dispersion of per-capita income across economies over time.

As for residential property, examples of studies employing β-convergence and σ-convergence include: Communities and Local Government (2010), Cook (2012), Nissan and Payne (2013) and Wood et al (2015).

The convergence methodology we employ is based on that developed by Phillips and Sul (2007). This is a relatively new approach towards investigating panel convergence, originally formulated to analyse economic transition behaviour. An appealing feature of the methodology is that it allows for different time paths to be taken by the constituent series, and so allows for individual heterogeneity by clustering panel *time-series* observations. Furthermore, the Phillips and Sul (2007) approach, described in Section 4, avoids much of the criticism levelled at time-series approaches, including issues with unit roots/stationarity aspects and cointegration, and has found wide application in many areas[3].

Representative studies in the house price literature employing the Phillips and Sul approach together with their key findings are outlined. Overall, it is found that, based on different data and samples, the results provide wide support for the fact that housing markets in the respective countries studied do not converge to a single long-term path, but rather form 'clubs'.

Apergis and Payne (2012) examined US house prices by state over the period 1975 to 2010 and identify three convergence clubs. They make the point that one aspect of the dynamics of regional prices implies that house prices converge as a result of the so-called 'ripple effect', where the transmission of shocks in regional house prices impacts on other regions. In this regard, they cite factors such as labour mobility and migration, equity conversion, spatial arbitrage and spatial patterns, Meen (1999), as being important determinants driving convergence.

Kim and Rous (2012) investigate house price convergence in a panel of US states and metropolitan areas and find evidence of multiple convergence clubs. They examine the general characteristics of the convergence clubs and the influential factors. They report that there exist a variety of different transitional (dynamic) patterns of prices across US states.

---

[3] Examples of studies applying the Phillips and Sul approach are provided in the *References* and include: Apergis et al (2011), Fritche and Kuzin (2011), Apergis and Gabrielsen (2012), Bartkowska and Riedl (2012), Montfort et al (2013), Apergis et al (2014), Borsi and Metiu (2015), Rughoo and You (2015), Antonakakis et al (2017) and Lyncker and Thoennessen (2017).

Their results indicate that housing supply regulation together with climate are important in determining converging club membership.

Montanes and Olmos (2013) also analyse US house prices, using monthly data over the period 2000:M1-2013M3, and find evidence for three clubs. Importantly, by allowing for sample changes they comment that the number and composition of clubs may alter over time.

Early work indicating a degree of segmentation across UK residential markets was demonstrated by Cook (2003) and Meen (2005). However, the paper by Montagnoli and Nagayasu (2015) appears to be the first paper to quantify the number of steady states in the UK housing market. Applying the Phillips and Sul methodology, they analyse convergence and spill-over effects of house prices in twelve UK regions over the period 1983-2012, finding that four groups of clusters exit. Their results complement previous studies looking at the UK housing markets, where there was no evidence of a single steady state. Furthermore, their findings support the ripple effect hypothesis, in that the London market influences residential prices in other regions.

Employing data over the period 1995-2017, evidence for house price convergence in 348 administrative locations in England and Wales was investigated by Holmes et al (2019). A large disaggregated panel comprising of a variety of types of housing (detached, semi-detached, terraced houses and flats) is used in the analysis. Interestingly, their results show that there are four convergence clubs for these four different categories of residential property.

Regional house price convergence in Spain and the factors driving convergence over the period 1997 to 2007 was investigated by Blanco et al (2016). In an interesting study they report that regional prices do not converge to a common trend, and identify four separate converging groups, indicating a degree of segmentation in the Spanish housing market. An ordered logit model is applied in order to account for the club membership, finding that provinces belonging to the same club had similar economic size, demographic composition by age and a similar number of vacant properties as a percentage of total stock of housing.

For Poland a few convergence studies exist. Gnat (2014) found beta, gamma and alpha convergence in the agglomeration area of Szczecin, implying that the satellite cities behave very similarly to Szczecin. Dittman (2014) using quarterly data and a sample extending over the period 2007–2012, analysed residential prices in 16 voivodeship capital cities using gamma-convergence, finding evidence only for weak convergence for the 11 smaller cities.

4. Phillips and Sul convergence methodology

As noted in Section 3, the methodology we use in looking to identify potential groups of cities where residential prices cluster was developed by Phillips and Sul (2007), and a brief overview of the approach follows.

The framework allows for the transition of the variable under investigation towards a long-run growth path *or* a common steady state equilibrium level. Furthermore, and importantly, as noted, the technique is robust to potential econometric violations of standard econometric assumptions, including stationarity issues. Originally, it was formulated in the context of convergence in economic growth for panels of countries, but has been applied much more widely as indicated earlier.

In addition to detecting the presence or absence of convergence, a key advantage of the so-called Phillips and Sul clustering *log-t* test algorithm is that it can reveal whether club formation is present. The test does not necessitate any specific assumptions, for example, the order of integration of the variables, and allows for cases where the individual time-series may be transitionally divergent. Indeed, the methodology enables the detection of convergence, where other methods such as stationarity tests fail. Stationary time-series methods are unable to detect the asymptotic co-movement of two time-series and therefore erroneously reject convergence.

The Phillips and Sul procedure is a semi-parametric clustering algorithm. A noteworthy feature of the analytical framework is that the rejection of the null hypothesis of convergence *does not* imply evidence against convergence at the level of *sub-groups* (clubs) within the panel. That is to say, it may be the case that within the full panel sub-groups may convergence i.e. there may be clusters, known as 'convergence clubs', within the panel.

The detailed presentation of the methodology developed by Phillips and Sul (2007, 2009) notes that the approach is essentially a nonlinear time-varying single factor model, that allows time-varying individual and common factor components. It recognises that common and idiosyncratic components are at work, providing a framework for modelling the transitional dynamics in panel models. The procedure is briefly outlined.

Given a panel of data, showing the residential price per square metre of city i at time t, represented as X_{it}, with N cities, with prices measured over the period t = 1…T. X_{it} is decomposed into two components, one systematic, g_{it}, and one transitory a_{it}

$$X_{it} = g_{it} + a_{it} \qquad (1)$$

Phillips and Sul (2007, 2009) transform (1) such that the common and idiosyncratic components in the panel are separated as follows:

$$X_{it} = \left(\frac{g_{it}+a_{it}}{\mu_t}\right)\mu_t = \delta_{it}\mu_t, \text{ for all } i,t \qquad (2)$$

This way the variable of interest, X_{it}, is decomposed into two components, where μ_t is a single common (trend) component for all cities and δ_{it} is a time varying idiosyncratic element, which measures the deviation of city price *i* from the common path, the trend component, defined by μ_t. It is assumed that *all* cities will converge, at some point in the future, to the steady state if $\lim_{k\to\infty} \delta_{it+k} = \delta_i = \delta$, for all *i* = 1…N.

It should be noted that the paths (mean reversion) to the steady state across cities may differ significantly. Phillips and Sul (2007, 2009) eliminate the common component μ_t by rescaling the panel average.

4.1 Testing for convergence: the *log-t* test

Phillips and Sul (2007) recommend a simple time-series regression test for convergence, the so-called *log-t* regression test for the null hypothesis of convergence. They propose two notions of convergence*: i) convergence in rates* or *relative* convergence, where the variables have the same rate of change (growth) and ii) convergence *in levels* or *absolute* convergence, where the variables converge to the same long-term value.

The average difference between δ_{it} and δ_i is assumed to decline over time at a rate proportional to $1/(t^{\alpha}\log(t+1))$, where α is a measure of the speed of convergence parameter. Given that there are two different types of convergence, Phillips and Sul show that *relative* convergence is ensured if $\alpha \geq 0$. Thus, the convergence hypothesis is:

$$H_0 : \delta_{it} = \delta \;\; \forall_i \; and \; \alpha \geq 0 \qquad (3)$$

while the alternative is, $H_A : \delta_{it} \neq \delta \;\; \forall_i \; or \; \alpha < 0$

For *absolute* converge, the null hypothesis is:

$$H_0 : \delta_{it} = \delta \;\; \forall_i \; and \; \alpha \geq 1$$

The null hypothesis can be tested by using equation (4):

$$log\left({}^{H_1}/_{H_t}\right) - 2logL(t) = c + blogt + \mu_t \qquad (4)$$

where $L(t) = \log(t+1)$, $H_t = \frac{1}{N}\sum_{i=1}^{N}(\hat{h}_{it} - 1)^2$ and $\hat{h}_{it} = \frac{\hat{X}_{it}}{\frac{1}{N}\sum_{i=1}^{N}\hat{X}_{it}}$ and $b = 2\alpha$.

Hypothesis (3) implies that, all other things equal, a large $log\left(H_1/_{H_t}\right)$ corresponds to a large b. Consequently, $H_t \rightarrow 0\ as\ t\ \rightarrow \infty$, suggesting that $h_{it} \rightarrow 1\ as\ t\ \rightarrow \infty$. The latter implies that X_{it} approaches the cross-sectional average, thus indicating evidence of convergence. Alternatively, a negative *b* provide*s* evidence of non-convergence. Thus, the convergence hypothesis is tested by the null hypothesis of *b = 0* against the alternative of non-convergence, *b < 0*, in equation (4). The test, which Phillips and Sul call the *log-t panel convergence test*, involves a one-sided t test of the inequality of the null hypothesis, using a heteroskedasticity and autocorrelation consistent standard error (HAC). It should be noted that under the null hypothesis in equation (3), the model allows for transitional periods, allowing for heterogeneity across cities in the short and medium-term. Hence, the long-run implication of equation (3) is that short-run deviation from the long-run may occur, however long-term trends bind the different city residential prices together.

4.2 Sub-groups of convergence clubs

As noted earlier, an appealing feature of the Phillips and Sul analytical framework is that the rejection of the null hypothesis of convergence *does not* imply evidence against convergence at the level of *sub–groups* of cities within the panel. That is to say, it may be the case that within the full panel sub–groups may convergence i.e. there may be clusters, known as *'convergence clubs'* within the panel (dual alternative hypotheses).

Consequently, a rejection of the null hypothesis for the panel as a whole does not necessarily imply that there is no convergence amongst sub-sets of cities, which did not form an initial convergence club consisting of all cities; there may exist multiple equilibria for different cities. In this regard, equations (5) and (6) present alternative hypotheses: equation (5) finds that cities diverge, however, in the alternative hypothesis, equation (6), some cities converge i.e. δ_{it} converges to δ_k. For K such groupings, G= [G_1, G_2, G_3, …, G_K], where the number of funds across counties sums to the total number N.

$$H_A : \delta_{it} \neq \delta\ or\ \alpha < 0 \qquad (5)$$

$$
H_A: \delta_{it} = \begin{cases} \delta_1, \ \alpha \geq 0 \ \ if \ i \in \ G_1 \\ \delta_2, \ \alpha \geq 0 \ \ if \ i \in \ G_2 \\ \quad\quad\quad\quad . \\ \quad\quad\quad\quad . \\ \quad\quad\quad\quad . \\ \delta_K, \ \alpha \geq 0 \ \ if \ i \in \ G_K \end{cases} \tag{6}
$$

Note: Phillips and Sul (2007,2009) provide details of a 4-step procedure for identifying such sub-groups.

Summarising, the combination of the null hypothesis, equation (3), and the dual alternative hypotheses, shown in (5) and (6), provide considerable flexibility with respect to the underlying theoretical framework and account for several possibilities. Essentially, the Phillips and Sul approach is a data driven algorithm investigating the possibility of the existence of convergence clubs (Du, 2017).

5. Phillips and Sul estimation results

Implementation considerations: As convergence is a long-term concept, Phillips and Sul (2007) recommend that the trend component is extracted from the series and that it is used for investigating convergence. Accordingly, we extract the trend using the Hodrick-Prescott filter (Hodrick and Prescott, 1980; Hodrick and Prescott 1997) and use this in the analysis. Employing quarterly data, we use a smoothing parameter of 1,600 (Ravn and Uhlig, 2002).

The estimate of the *log-t* regression (equation 4) proposed by Phillips and Sul for the whole panel of 16 cities, resulted in a $\hat{\beta}$ value of -0.663 with a t-statistic of -15.704. This rejects the null hypothesis of convergence amongst the sixteen cities, as the t-value fell well below the critical value of -1.65 (5% significance level), indicating divergence within the full panel. Hence, the Phillips and Sul convergence tests clearly makes evident that Polish regional cities residential prices do not converge to a single price level. However, as discussed, rejection of the null of convergence for the full group does not mean that each city follows its own independent path, as sub-groups may be present which do converge.

We next apply the Phillips and Sul clustering algorithm outlined in Section 4.2 to examine the presence of any sub-groups (clubs) of cities. Table 1 shows the results of applying the *log-t* test to the (log) of residential prices across the 16 cities.

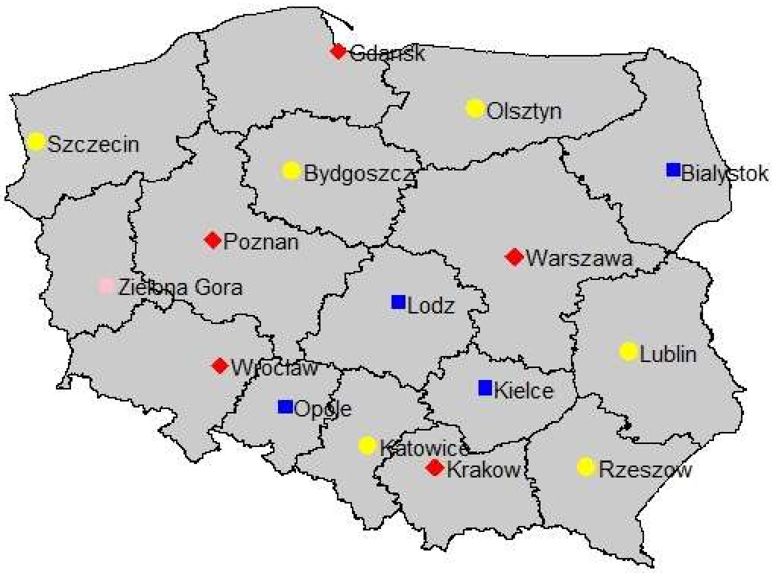
Table 1: Classification of club convergence 2007:Q1-2018:Q4

Club	Cities	$\hat{\beta}$	t-stat	$\hat{\alpha}$
1	5	0.171	2.725	0.085
2	6	2.019	5.845	1.009
3	4	0.699	0.928	0.349

Note: Applied truncation parameter r = 0.3 based on 48 observations; applied critical value c=0; t-statistic at the 5% significance level = -1.645; $\hat{\alpha}$ = speed of convergence; number of diverging cities = 1 (Zielona Gora); estimation undertaken with R.

The results show that there are 3 clubs. Furthermore, one city, Zielona Gora, was reported as being a divergent city, is not a member of any club and is therefore excluded from further analysis. The hypothesis of convergence for the cities within each of the three clubs cannot be rejected, as the $\hat{\beta}$ t-statistics are all greater than -1.65 (5% significance level). The parameter determining the speed of convergence is α, and Table 1 shows that the convergence speeds (the $\hat{\alpha}$ values) are positive, but differ for the three clubs. The middle-income cities convergence rate, club 2, is the highest of the three groups. Club 2 has an $\hat{\alpha}$ value greater than 1, indicating that the tendency is for prices to converge *in absolute terms.* The $\hat{\alpha}$ value for clubs 1 and 3 are positive and less than 1, indicating that cities in these two clubs neither diverge from nor converge to the same level of price, but converge conditionally, that is, prices change *at the same rate* i.e. growth rates are similar within each club. Furthermore, following Phillips and Sul (2009), as the $\hat{\alpha}$ value of club 3 is not statistically different from zero, this is taken as evidence that club 3 displays weaker convergence compared to the other two clubs. The geographic distribution of the clubs is shown in Figure 3.

Figure 3: Distribution of City clubs



5.1 Residential price characteristics of the clubs

We summarise the price characteristics of the three clubs in Table 2. For club 3 (the largest and richest cities) in 2007, initial residential price levels were in excess of 60% greater than those for clubs 1 and 2, where initial levels are broadly similar. Comparing the average residential prices for the clubs in 2018, club 3 has the highest residential prices, by far exceeding those for clubs 1 and 2, clubs 1 and 2 having broadly similar prices. However, the gap between club 3 and clubs 1 and 2 has narrowed over the period.

Average prices for each club in 2007 are shown. Over the period 2007-2018, clubs 1 and 2, experienced the highest average annual growth rates, the distribution of growth rates within each club is also shown. Club 3 had, by a considerable margin, the highest average prices in 2007 and the lowest average annual increase in prices over the period 2007-2018. Indeed, within all three clubs (excluding Gdansk in club 3, which had a growth rate of 3.16%), there was an inverse relationship between initial club prices in 2007 and subsequent growth in prices. (Interestingly, this has parallels in the income per capita growth literature, where initial *structural income* conditions are inversely related to subsequent growth in income, so-called Beta-convergence.)

Table 2: Summary price statistics, 2007-2018

Club	Average Price 2007	Average Price 2018	Average annual growth rate (%)
1	4281	5073	3.42 [2.32 to 4.09]
2	4157	5419	4.41 [3.55 to 5.06]
3	6922	7233	0.97 [-0.19 to 3.16]

Note: Average Prices for 2007/2008 are PLN/m^2 for the year.
Average annual growth rates are averaged across individual cities for each club.

5.2 Temporal evolution of price trends: transition paths

The Phillips and Sul approach allows the transition path of each city's prices to be traced out. This enables the shape of the paths to be compared, showing where there is transitional divergence followed by a catching-up process, allowing prices to converge.

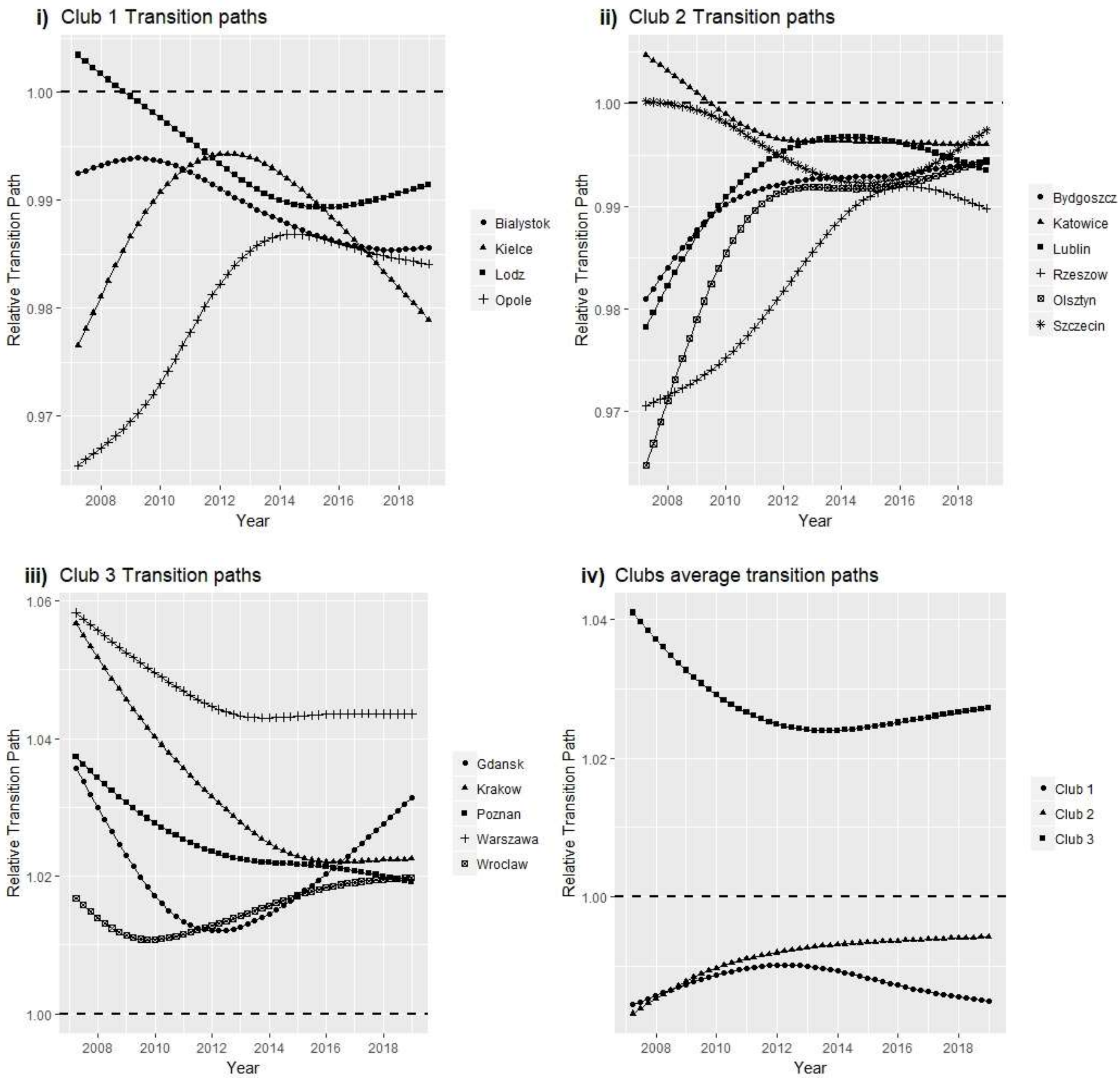
In order to examine the transition paths of prices for the clubs, the relative transition path of each city's residential (log) prices over time are profiled in Figure 4, panels i), ii) and iii). As shown in Section 4.1, the relative transition paths are determined by the transition coefficient $\hat{h}_{it} = \frac{\hat{X}_{it}}{\frac{1}{N}\sum_{i=1}^{N}\hat{X}_{it}}$ over time, where the $\hat{X}_{it}$ variables represent the individual cities log of prices over time and the denominator is the average of the whole panel's log5prices over time.

In the convergence club literature, notions of convergence reflect initial (starting) conditions. The cities in each club had differing initial conditions, that is, different initial residential prices in 2007, with clubs 1 and 2 having lower initial prices than those in club 3 (Table 2). This is, largely, reflected in the profiles of the relative transition paths for the cities in clubs 1 and 2, being quite different from those in club 3. Three of the four cities in club 1 and four of the 6 six cities in club 2 initially display a catching-up (relative rising prices) tendency, where prices tend towards the average of all cities (the value 1). By way of contrast, all five cities in club 3 display a fall-off in prices relative to the panel average of all cities. Relative to the average of all cities, the cities in club 3 display a profile of higher relative prices, all having relative values in excess of 1. Panel iv) summarises the average position for all of the cities within each club.

The transition paths reflect the long-term process of divergence from the panel average and the subsequent catch-up to the average, conditional on the dynamics of the underlying factors driving residential prices within each city.

Clearly, as shown, over time cities *within each club* can behave differently, even diverging at times, but with an overall tendency towards convergence within each club.

Figure 4: Relative transition paths for three convergence clubs



6. Ordered logit methodology

Whilst the Phillips and Sul *log-t* convergence test approach can identify convergence clubs, it does not address how such clubs are formed. The formation of clubs may indicate that common factors amongst cities lead them to converge to a similar steady state. Consequently, the question arises, what are the underlying factors which drive club formation and how do they influence the likelihood of observing a city in a given club? We are looking to identify the factors driving club membership and quantify the effects of each factor for each club.

The clubs can be regarded as an ordinal variable which can be classified as being ordered in some meaningful way (Section 6.1). The dependent variable represents the club to which the city belongs. It is possible to order and rank the three clubs determined by the Phillips and Sul procedure, and so we employ an ordered logit model in order to assess the potential drivers determining the clubs, examples being: Bartkowska and Reidl (2012), Kim and Rous (2012), Lyncker and Thoennessen (2017), Blanco et al (2016) and Zhang et al (2019).

There are a variety of ordered logit models which can be employed, as discussed in Greene (2012), Agresti (2015) and Fullerton and Xu (2016), and we employ a *proportional odds model*. Conditional on the variables driving city residential prices, the model estimates the probabilities of a city belonging to the various clubs. Based on the highest probability, a city is allocated to one of the three clubs.

Let C denote the convergence club which can take on values {1,2,…,J}, where J is the number of convergence clubs, being three in the current context of the Polish cities. Let x be a vector of conditioning variables which drive club membership.

The ordered logit model can be formulated as a latent variable model, namely:

$$y'_i = \boldsymbol{x}\beta + \epsilon_i.$$

$$\text{where, } y_i = j \;\; if \;\; \tau_{j-1} \leq y'_i < \tau_j \,, \;\; j = 1 \ldots J$$

y'_i is the latent variable ranging from ∞ to $-\infty$, and ϵ_i is a random error, $\boldsymbol{x}$ is the vector of independent variables, βs are the slope coefficients, τ_j are the thresholds and J is the number of categories(clubs). The thresholds τ_1 through $\tau_{J-1}$are the parameters to be estimated, assuming that $\tau_0 = -\infty$ and $\tau_J = \infty$.

Specifically, the latent variable y'_i can be regarded as the tendency of a city to belong to one of the three clubs, and the threshold (transition) parameters τ_j separate the clubs.

Consequently, when the underlying latent variable crosses a threshold τ_j, the city club membership changes.

The ordered logit model is defined as the probability model,

$$ln\left(\frac{Pr(y\le j|\boldsymbol{x})}{Pr(y>j|\boldsymbol{x})}\right) = \tau_j - \boldsymbol{x}\beta, \quad j = 1, \dots, J-1 \tag{7}$$

where $ln\left(\frac{Pr(y\le j|\boldsymbol{x})}{Pr(y>j|\boldsymbol{x})}\right)$ is known as the *log-odds*.

The probabilities of belonging to a given club are evaluated at the mean values of the $\boldsymbol{x}$ variables and are determined as follows:

$$Pr(y = 1|\boldsymbol{x}) = \frac{exp(\tau_1 - \boldsymbol{x}\beta)}{1 + exp(\tau_1 - \boldsymbol{x}\beta)} \quad for\ j = 1$$

$$Pr(y = j|\boldsymbol{x}) = \frac{exp(\tau_j - \boldsymbol{x}\beta)}{1 + exp(\tau_j - \boldsymbol{x}\beta)} - \frac{exp(\tau_{j-1} - \boldsymbol{x}\beta)}{1 + exp(\tau_{j-1} - \boldsymbol{x}\beta)} \quad for\ j = 2, \dots, J-1$$

$$Pr(y = J|\boldsymbol{x}) = 1 - \frac{\exp(\tau_{J-1} - \boldsymbol{x}\beta)}{1 + \exp(\tau_{J-1} - \boldsymbol{x}\beta)} \quad for\ J$$

In order to assess the importance of variables in determining club membership, we estimate the *marginal effects* of the predicted probabilities. The marginal effects provide an estimate of how the probability of belonging to a club changes for a unit change in one of explanatory variable, holding all other variables fixed at their sample averages[4].

6.1 Factors shaping club membership

We borrow concepts used in identifying convergence clubs from the economic/GDP per capita income growth literature, for example Galor (1996), Phillips and Sul (2007, 2009), Fritche and Kuzin (2011) and Montfort et al (2013). The convergence hypothesis postulates that both initial (starting) conditions *and* 'structural' characteristics matter. In the context of Polish residential prices, 'initial' conditions, such as levels of prices in 2007 (the start year of the analysis) are likely to be a factor in helping to explain which club a city will belong to. Secondly, market characteristics will shape the trajectories of club convergence. So, both

[4] Agresti and Tarantola (2018).

initial conditions and evolving market characteristics are likely to influence club formation. Consequently, we employ both in our analysis.

Several potential variables suggest themselves as representing initial conditions in 2007 including[5]: population, housing stock, housing per 1000 of population and average residential prices. Structural characteristics in city residential property markets include variables such as: cost of debt, population/population growth, incomes/income growth, housing stock/housing stock growth, new construction and unemployment rates. Structural characteristics will determine the long-run growth path of residential property prices.

Table 3 provides summary measures of some of the key variables considered as potential factors in accounting for club membership.

Table 3: Club features summary statistics

Club	Averages 2007-2018			Housing stock per 1000
	Population (000's)	Unemployment rate (%)	Monthly wage (PLN)	
1	346	8.67	3398	426
2	296	6.83	3417	413
3	823	4.05	3888	466

The average population for club 3 is far greater than that for clubs 1 and 2 and the unemployment rate is also the lowest. Similarly, average monthly wages are the highest for club 3, as is the housing stock per 1000 of population.

As for ranking, the three clubs can be ranked on the basis of 2018 average prices. As shown in Table 2, club 3 has the highest average prices, followed by club 2 and finally club 1. Consequently, the clubs can be ranked in a logical way, according to the level of residential prices, from lowest to highest residential prices. As there are three clubs, the dependent variable, y, takes on values from 1 to 3 and, in combination with the driver variables, the ordered logit model is used to allocate the cities to one of the clubs.

6.2 Ordered logit model estimation results

We undertook a top-down estimation approach in arriving at a parsimonious model involving a sub-set of potential variables. Given the relatively small sample, the potential for testing the

[5] Prices and interest rates originate from the Narodowy Bank Polski (NBP, 2019). The remaining variables are from the Local Data Bank of the Central Statistical Office of Poland (CSO, 2019).

robustness of any derived model was restricted. However, the model presented below satisfied a priori, statistical and performance criteria.

The results of applying the ordered logit model, the parameter estimates together with their attendant robust standard errors are shown in Table 4. The estimated coefficients in logistic models are reported in log-odds form, equation (7).

Table 4: Estimated ordinal logit model parameters

Variable	**Coefficient**	**Odds Ratio ($exp^{-\beta}$)**
Price 2007	0.002*** (0.001)	1.000
Population growth	3.313*** (0.320)	0.036
Housing stock growth	-0.847*** (0.154)	2.332
Unemployment rate	-1.021** (0.549)	2.776
Wage growth	1.305* (0.916)	0.271
Transition parameter 1|2	6.523*** (0.180)	-
Transition parameter 2|3	13.851*** (3.490)	-

*Note: One-tail test. *p<0.1; **p<0.05; ***p<0.01.*
Growth variables are annual % averages over 2007-2018.

We first test for the parallel regression assumption, which was assumed in the method of estimation. The assumption of the validity of proportional odds (equal coefficients across the regimes) needs to be tested as it is frequently invalid, and so we employ Brant's test (Brant, 1990) to test this assumption. The test does not reject the null for either the regression or at the individual coefficient level.

All of the estimated coefficients had the a priori expected signs. All of the variables, apart from wage growth, are significant at either the 1% or 5% significance level, wage growth being significant at the 10% level. The coefficients in an ordered logit model are interpreted as follow. The coefficient values have the usual interpretation; if $\beta_k > 0$ then a 'positive' effect, and vice versa. A positive coefficient indicates an increased chance that the city with a higher value of the independent variable, x_k, will be in a higher club, whist a negative coefficient suggests that the city with a higher value of the variable x_k will be in a lower valued club.

Also reported in Table 4 are the corresponding odds ratios, estimated for each coefficient as $exp^{-\beta_i}$, where β_i is the variable's coefficient. The odds ratios allow for easier interpretation, and are interpreted as the impact (increase or decrease) in the odds of being in a different club, given a unit increase in the $x_k$variable of interest. In the current context, values of less than 1 make it less likely of being in a lower club and values greater than 1 more likely of being in a lower club. The reported transition parameters are used to calculate the probability of a city being in a particular club (Section 6), the reported values being the 'boundaries' between the clubs (see Agresti 2015).

Based on their respective highest probability values, each city is allocated to one of the three clubs. In 13 out of 15 cases, the cities were correctly classified as members of the club which resulted from the Phillips and Sul classification procedure.

To enable the interpretation of the impact of the variables, Table 5 reports the marginal effects, computed at the mean value of the variables. These show the impact of each variable on the *change in probability* of belonging to a club, given a small change in the variable, effectively showing the estimated change in probability of staying in a club.

Table 5: Marginal effects

		Marginal effects (probabilities)	
Variable	**Club 1**	**Club 2**	**Club 3**
Price 2007	-0.00003 (-0.0048)*	-0.00011 (-0.0403)*	0.00014 (0.0451)*
Population growth	-0.0335	-0.148	0.1817
Housing stock growth	0.00881	0.0377	-0.04651
Unemployment rate	0.01061	0.04542	-0.05603
Wage growth	-0.01356	-0.05807	0.07163

Note: Growth figures represent annual % averages over 2007-2018.
* Based on 2007 average prices being 5% higher (see the following comment).

The *Price 2007* variable, the initial conditions variable as discussed earlier, warrants commenting on. The reported coefficient, 0.0002 in Table 3, represents the impact of a unit increase of 1(PLN) and is a small figure compared to average prices in 2007 (see Table 1). Being positive, the higher initial residential price increases the probability of a city belonging to a higher price club by a very small amount. The change in marginal probabilities, the increase in probability of being in a higher club (club 3) is positive, and the probability of being in a lower club (clubs 1 and 2) is reduced, are very small changes in probability.

Consequently, in this situation a unit change in the initial price will have next to no impact in increasing the probability of a city being in a different club. In order to have a more realistic assessment of the impact higher initial prices, say prices being 5% higher, the resulting change in probability of the city being in a higher club would be more marked, which is shown in parentheses for the Price 2007 variable. Clearly, the marginal probabilities reflect a more meaningful impact of a city being in a higher club thank is currently the case.

Changes in population growth appear to have the largest impact. An increase of 1% in annual growth increase the probability of being in club 3 by 18.2 percentage points, with a 14.8 percentage point decrease in probability of being in club 2, and a 3.35 percentage decrease in probability of belonging to club 1. Housing stock growth increases the probability of being in a lower club by 4.65 percentage points for club 3, with positive increased probabilities of membership of clubs 1 and 2. The marginal impact of a wage growth increase of 1% increases the probability of club 3 membership by 7.2 per centage points, whilst decreasing the probability of club 1 and club 2 membership by 1.4 and 5.8 percentage points respectively.

7. Conclusions and observations

The research has shown the existence of three residential price convergence groups in fifteen key Polish city markets, tending to converge towards respective common club price levels. Interestingly, those clubs are not clustered geographically, even though one could expect that cities which are regionally close show similar price behaviour.

Although the determination of clubs is (endogenously) data determined and avoids pre-specifying potentially arbitrary a priori groupings, it is nonetheless atheoretical. In order to have an understanding of why the residential markets group the way they do further insight is required.

Club convergence hypothesis studies investigating economic or per capita income growth, demonstrate that identical ‘structural characteristics’ in different economies/regions lead to convergence in economic/income levels in the long-run, provided there are similar *initial* conditions for each club. In looking to establish the factors which determine the three clubs, we incorporate both initial and structural conditions in the analysis.

To explain the existence of the clubs, an ordered logit model demonstrated that a number of supply and demand variables, together with initial price conditions, accounted for the three club groupings. Indeed, 13 of the 15 cities investigated were correctly classified as belonging to the clubs obtained by the Phillips and Sul procedure.

Whist the reported results provide an insight into the city groupings, the formation of residential clubs will likely be determined by other factors which have not been taken into account in our analysis. A more detailed analysis and profile of how cities transition from one club to another and the evolution of clubs over time would provide insights into the dynamics of clubs. Furthermore, it would be insightful to explore the interaction and links between prime markets and secondary residential property markets.

Considering club convergence from a broader perspective, we would comment that the question of the existence of residential clubs is important for several reasons. The existence of clubs provides a better understanding of the functioning of residential markets, rather than just merely looking at each city market in isolation. This will assist in more insightful modelling, and forecasting, of heterogeneous residential market groupings. It will help in understanding the extent to which groups of cities display differences in the behaviour of residential prices.

The evolution of the housing markets can have a significant influence on local economic activity, on the affordability of housing, on relocation costs and on labour mobility between regions. Convergence and divergence in residential prices at different points in the property cycle will have implications for economic policy. Consequently, the identification and segmentation of regions/cities will enable targeted policy by government and/or local authorities, where the impact of policy is likely to be most effective.

We would remark that further research could usefully investigate the sensitivity of the results to starting and ending dates, together with the dynamics of how cities might transfer from one club to another.

Finally, from an investment perspective, many funds now invest in residential property markets, and consequently, there are risk reduction/diversification implications. Investing in converging markets may restrict risk control opportunities as such markets will be ‘similar’ Yunus and Swanson (2013). However, an understanding of how, and which, residential markets converge also presents investment opportunities for funds.

References:

Agresti A (2015) *Categorical Data Analysis*, Third Edition. John Wiley & Sons, Inc.

Agresti A and Tarantola C (2018) Simple ways to interpret effects in modelling ordinal categorical data, *Statistica Neelandica*: pp 1-14. https://doi.org/10.1111/stan.12130.

Antonakakis N, Christou C, Cunado J and Gupta R (2017) Convergence Patterns in Soverign Bond Yields Spreads: Evidence from the Euro Area. *Journal of International Financial Markets, Institutions & Money*: pp 1-30. doi:http//dx.doi.org/10.1016/j.intfin.2017.03.002.

Apergis T, Christou C and Payne J (2011) Political and Institutional Factors in the Convergence of International Equity Markets: Evidence from the Club Convergence and Clustering Procedure. *Atlantic Economic Journal* 39(1): 7-18.

Apergis N and Gabrielsen A (2012) Testing for Convergence in the European Insurance Sector: A Non-Linear Factor Approach. *Asian Journal of Finance & Accounting* 4(2):1-17.

Apergis N and Payne J (2012) Convergence in U.S. house prices by state: evidence from the club convergence and clustering procedure. *Letters in Spatial and Resource Sciences* 5(2): 103-111.

Apergis N, Christou C and Miller SM (2014) Country and industry convergence of equity markets: International evidence from club convergence and clustering. *North American Journal of Economics and Finance* 29: 36-58.

Augustyniak H, Łaszek J, Olszewski K and Waszczuk J (2014) Modelling of cycles in the residential real estate market. *Kwartalnik Nauk o Przedsiębiorstwie* 2: 63-77.

Baldowska G, Leszczynski R and Myszkowska B (2014) Convergence and differentiation processes in local markets and structural changes (comparison of 16 markets in Poland). *National Bank of Poland Working Paper No. 174*. Economic Institute, Warsaw.

Bartkowska M and Riedl A (2012) Regional convergence clubs in Europe: identification and conditioning factors. *Economic Modelling* 29(1): 22–31.

Barro RJ (1991) Economic Growth in a Cross Section of Countries. *Quarterly Journal of Economics* 106(2): 407-443.

Barro RJ and Sala-i-Martin X (1992) Convergence. *Journal of Political Economics* 100(2): 223-251.

Baumol WJ (1986) Productivity Growth, Convergence, and Welfare: What is the Long-Run Data Show. *American Economic Review* 76(5): 1071-1085.

Blanco F, Martin V and Vazquez G (2016) Regional House price convergence in Spain during the housing boom. *Urban Studies* 53(4): 775-798.

Borsi MT and Metiu N (2015) The Evolution of Economic Convergence in the European Union. *Empirical Economics* 48(2): 657–681.

Brant R (1990) Assessing Proportionality in the Proportional Odds Model for Ordinal Logistic Regression. *Biometrics* 46(4): 1171-1178.

Ciarlone A (2015) House price cycles in emerging economies. *Studies in Economics and Finance* 32(1): 17–52.

Cook S (2003) The convergence of regional house prices in the UK. *Urban Studies* 40(11): 2285-2294.

Cook S (2012) *β*-convergence and the cyclical dynamics of UK regional house prices. *Urban Studies* 49(1): 203-2018.

Communities and Local Government. (2010), Housing and Regional Economic Disparities: economics paper 5. *DCLG economics papers*.

CSO (2019) Local Data Bank. *Central Statistical Office of Poland*. https://bdl.stat.gov.pl/BDL/start.

Dittmann I (2014) Gamma konwergencja cen na lokalnych rynkach mieszkaniowych w Polsce. *Studia Ekonomiczne*, *Zeszyty Naukowe Uniwersytetu Ekonomicznego w Katowicach* 181: 195-207.

Du KR (2017) Econometric convergence test and club clustering using Stata. *Stata Journal*17(4): 882-900.

Eurostat (2018) Living conditions in Europe. *Chapter 3 Housing quality*. European Union, 2018.

Fritche U and Kuzin V (2011) Analysing convergence in Europe using the non-linear single factor model. *Empirical Economics* 41(2): 343-369.

Fullerton AS and Xu J (2016) *Ordered Regression Models Parallel, Partial, and Non-Parallel Alternatives*. CRC Press:Taylor & Francis Group, LLC.

Galor O (1996) Convergence? Inferences from theoretical models. *Economic Journal* 106(437): 1056–1069.

Gnat S (2014) Konwergencja cenowa na lokalnych rynkach nieruchomości w aglomeracji szczecińskiej w latach 2008-2013. *Studia i Prace WNEIZ US* 36(1): 213-225.

Greene WH (2012*) Econometric Analysis*, Seventh Edition. Pearson Education Limited, Harlow, England.

Hodrick RJ and Prescott EC (1997) Postwar U.S. business cycles: an empirical investigation. *Journal of Money,Credit and Banking* 29(1):1-16.

Hodrick RJ and Prescott EC (1980) Postwar U.S. business cycles: An empirical investigation. *Discussion Paper No. 451*. Carnegie-Mellon University.

Holmes JM, Otero J and Panagiotidis T (2019) Property heterogeneity and convergence club formation among local house prices. *Journal of Housing Economics* 43: 1-13. doi:https://doi.org/10.1016/j.jhe.2018.09.02.

Kim YS and Rous JJ (2012) House price convergence: Evidence from US state and metropolitan area panels. *Journal of Housing Economies* 21: 159-186.
Łaszek J, Augustyniak H and Olszewski K (2016a) FX mortgages, housing boom and financial stability–a case study for Poland (2005-2015). *The Narodowy Bank Polski Workshop: Recent Trends in the Real Estate Market and its Analysis*, 2015 Edition.

Łaszek J, Olszewski K and Waszczuk J (2016b) Monopolistic Competition and Price Discrimination as a Development Company Strategy in the Primary Housing Market. *Critical Housing Analysis* 3(2): 1-12.

Łaszek J, Olszewski K and Waszczuk J (2017) Regional Diversification of Saturation of the Housing Stock: Poland against the Background of Europe. Barometr Regionalny*, Analizy i Prognozy* 15(2): 7-17.

Leszczynski R and Olszewski K (2017) An analysis of the primary and secondary housing market in Poland: evidence from the 17 largest cities. *Baltic Journal of Economics* 17(2): 136-151.
Lyncker K and Thoennessen R (2017) Regional club convergence in the EU: evidence from a panel data analysis. *Empirical Economics* 52(2): 525-533.

Mankiw NG, Romer R and Weil DW (1992) A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics* 107(2): 407-437.
Meen G (1999) Regional house price prices and ripple effect: a new interpretation. *Housing Studies*14(6): 733-753.

Meen G (2005) On the economics of the Barker Review of Housing Supply. Housing Studies, 20(6), pp 949-947.

Montagnoli, A. and Nagayasu, J (2015) UK house price convergence clubs and spillovers. *Journal of Housing Economics* 30(C): 50-58.

Montfort M, Cuestas KC and Ordonez J (2013) Real convergence in Europe: A cluster analysis. *Economic Modelling* 33(C): 689-694.

Montanes A and Olmos L (2013) Convergence in US house prices. *Economics Letters* 121(2): 152-155.
NBP (2019) Information on home prices and the situation in the housing and commercial real estate market in Poland in 2018 Q4. *Narodowy Bank Polski*.

Nissan E and Payne JE (2013) A Simple Test of σ-convergence in U.S. Housing Prices across BEA Regions. *Journal of Regional Analysis & Policy* 43(2): 178-185.

Phillips PC and Sul D (2007) Transition modelling and econometric convergence tests. *Econometrica* 75(6): 1771–1855.

Phillips PC and Sul D (2009) Economic transition and growth. *Journal of Applied Econometrics* 24(7):1153–1185.

Ravn MO and Uhlig H (2002) On adjusting the Hodrick-Prescott filter for frequency of observations. *Review of Economics and Statistics* 84(2): 371-376.

Rughoo A and You K (2015) Asian Financial Integration: Global or Regional? Evidence from Money and Bond Markets. *International Review of Financial Analysis* 48(C): 419-434.

Solow RM (1956) A contribution to the theory of economic growth. *Quarterly Journal of Economics* 70(1): 65-94.

Wood G, Sommervoll DE and de Silva A (2015) Do Urban House Prices Converge? *Urban Policy and Research* 34(2): 102-115.

Yunus N and Swanson PE (2013) A closer look at the U.S. housing market: modelling relationships among regions. *Real Estate Economics* 41(3): 542–568.

Zhang W, Xu W and Wang X (2019) Regional convergence clubs in China: identification and conditioning factors. *The Annals of Regional Science* 62(2): 327-350.